Cell PRESS
Structure
Article

Locked Tether Formation by Cooperative Folding of Rna14p Monkeytail and Rna15p Hinge Domains in the Yeast CF IA Complex

Maria Moreno-Morcillo,[1,2,3] Lionel Minvielle-Sébastia,[1,2] Sébastien Fribourg,[1,2] and Cameron D. Mackereth[1,2,*]
[1]Institut Européen de Chimie et Biologie and INSERM U869, 33607 Pessac Cedex, France
[2]Université de Bordeaux, 33076 Bordeaux Cedex, France
[3]Universidad de Zaragoza, Pedro Cerbuna 12, 50009, Spain
*Correspondence: c.mackereth@iecb.u-bordeaux.fr
DOI 10.1016/j.str.2011.02.003

SUMMARY

The removal of the 3′ region of pre-mRNA followed by polyadenylation is a key step in mRNA maturation. In the yeast *Saccharomyces cerevisiae*, one component of the processing machinery is the cleavage/polyadenylation factor IA (CF IA) complex, composed of four proteins (Clp1p, Pcf11p, Rna14p, Rna15p) that recognize RNA sequences adjacent to the cleavage site and recruit additional processing factors. To gain insight into the molecular architecture of CF IA we solved the solution structure of the heterodimer composed of the interacting regions between Rna14p and Rna15p. The C-terminal monkeytail domain from Rna14p and the hinge region from Rna15p display a coupled binding and folding mechanism, where both peptides are initially disordered. Mutants with destabilized monkeytail-hinge interactions prevent association of Rna15p within CF IA. Conservation of interdomain residues reveals that the structural tethering is preserved in the homologous mammalian cleavage stimulation factor (CstF)-77 and CstF-64 proteins of the CstF complex.

INTRODUCTION

The production of eukaryotic mRNA requires several processing events to form the mature molecule from the initial RNA polymerase II pre-mRNA transcript. One set of reactions includes cleavage of the pre-mRNA at a specific site in the 3′-untranslated region followed by polyadenylation (Mandel et al., 2008; Millevoi and Vagner, 2010; and references therein). Although the cleavage and polyadenylation processes can be uncoupled in vitro, these two steps are closely connected within the cell through shared protein factors and timing. Pre-mRNA 3′-end processing is intimately linked with transcription termination, and the resulting poly(A) tail is required for mRNA stability, efficient export to the cytoplasm, translation regulation, and post-transcriptional control. Fidelity in 3′-end processing is critical as primary components in yeast are essential, and deregulation of this process in humans can lead to disease (Danckwardt et al., 2008). Polyadenylation can also occur at alternative cleavage sites, making possible a higher level of gene transcript control (Lutz, 2008).

Within yeast, the processing machinery can be biochemically and functionally divided into the cleavage factor I (CF I) and the larger cleavage and polyadenylation factor (CPF) (reviewed in Mandel et al., 2008). The five proteins that make up CF I can be further separated into two purified factors: CF IA, which contains Rna14p, Rna15p, Pcf11p, and Clp1p, and CF IB, known as Hrp1p/Nab4p. The most related factor to CF IA in metazoa is the cleavage stimulation factor (CstF) complex consisting of CstF-77 (orthologous to Rna14p), CstF-64 (orthologous to Rna15p) and CstF-50. The mammalian orthologs of Clp1p and Pcf11p form an independent complex termed CFIIm (de Vries et al., 2000). There is no clear metazoan counterpart to Hrp1p.

Rna14p and Rna15p were the first two components identified from CF IA (Minvielle-Sebastia et al., 1991, 1994). Mutants in these genes exhibit long extended read-through transcripts lacking poly(A) tails that are rapidly degraded in the nucleus (Minvielle-Sebastia et al., 1991, 1994; Torchet et al., 2002). The aberrant transcripts are primarily due to defects in cleavage/polyadenylation, nonspecific poly(A) site selection and inefficient transcription termination (Mandart and Parker, 1995; Birse et al., 1998). A direct association between Rna14p and Rna15p has been demonstrated using yeast proteomics (Gavin et al., 2002; Nedea et al., 2003) and numerous biochemical and genetic studies (Noble et al., 2004; Gross and Moore, 2001b; Qu et al., 2007; Kessler et al., 1996; Legrand et al., 2007; Minvielle-Sebastia et al., 1991, 1994). Both proteins are found in the nucleus (Bonneaud et al., 1994), although Rna14p also appears to have a separate role within the mitochondria (Bonneaud et al., 1994; Rouillard et al., 2000). In addition, it has been shown that *rna14* and *rna15* mutant extracts cannot biochemically complement each other in pre-mRNA 3′-end processing assays (Minvielle-Sebastia et al., 1994), suggesting an uncommonly strong interaction that is resistant to component exchange to reconstitute a new active factor.

Rna14p is primarily composed of a large domain of HAT repeats (Preker and Keller, 1998), for which the structure has been determined for the homologous region in CstF-77 (Bai et al., 2007; Legrand et al., 2007). The atomic details explain the homodimeric nature previously described for Rna14p (Noble et al., 2004). The C-terminal region shares limited sequence similarity among the yeast and mammalian orthologs. However,

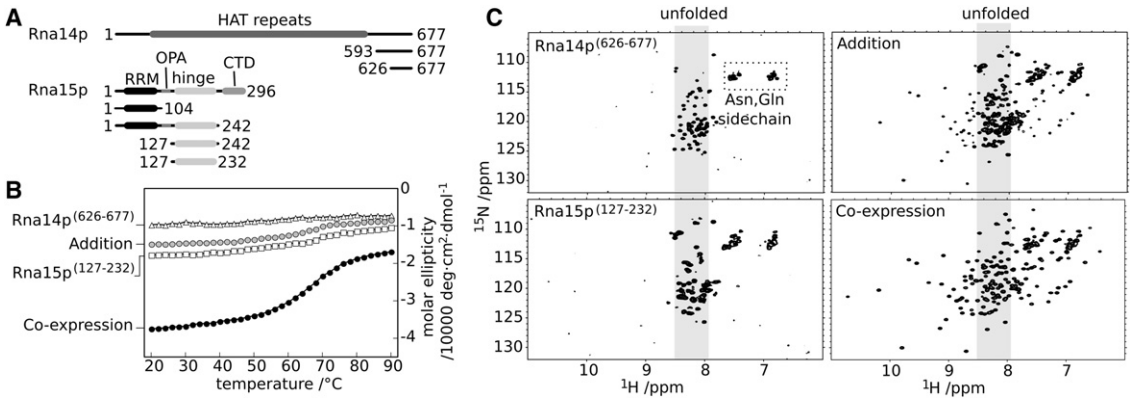

Figure 1. Identification of the Regions in Rna14p and Rna15p Responsible for Heterodimerization

(A) Schematic representation of truncation constructs, with initial interaction regions based on results from Legrand et al. (2007) and Qu et al. (2007). Association determined by bacterial coexpression is depicted in Figure S1.

(B) Thermal denaturation followed by circular dichroism (CD) spectroscopy at 222 nm of 15 μM His_6-Rna14p$^{(626–677)}$ (open triangles), His_6-Rna15p$^{(127–232)}$ (open squares), an equimolar addition of His_6-Rna14p$^{(626–677)}$ to His_6-Rna15p$^{(127–232)}$ (gray circles), and a coexpressed and His_6-tag removed sample of Rna14p$^{(626–677)}$/Rna15p$^{(127–232)}$ (black circles). The buffer contained 20 mM sodium phosphate (pH 6.5) and 50 mM NaCl.

(C) Two-dimensional $^1H^{15}N$-HSQC spectra at 303 K and 700 MHz corresponding to the same constructs as in (B). Samples were tested in 20 mM sodium phosphate (pH 6.5), 50 mM NaCl, 2 mM dithiothreitol, and 10% (v/v) ^2H_2O with protein concentrations of 200 μM except for the equimolar combination that contained 100 μM each of His_6-Rna14p$^{(626–677)}$ and His_6-Rna15p$^{(127–232)}$.

temperature sensitive mutations in this region in yeast (Minvielle-Sebastia et al., 1991), as well as in *Drosophila* (Simonelig et al., 1996), have indicated its importance.

Rna15p is the only component of CF IA that directly contacts the pre-mRNA (Kessler et al., 1996; Noble et al., 2004). The RNA recognition motif (RRM) domain from Rna15p has been structurally characterized bound to RNA (Leeper et al., 2010, Pancevac et al., 2010). The C-terminal domain of Rna15p is also highly conserved, and structure determination of the corresponding mammalian domain reveals a three-helix bundle that is required for interaction with Pcf11p and for 3′-end processing but not for transcription termination (Qu et al., 2007). The middle region of Rna15p shares homology with the central hinge domain of CstF-64, and has been implicated in the association between Rna15p/CstF-64 and Rna14p/CstF-77 (Hockert et al., 2010; Legrand et al., 2007; Qu et al., 2007). Despite the importance of this interaction revealed through genetics and biochemistry, there is no description of the molecular mechanism by which the two proteins associate.

To shed light on the atomic details that create the strong interaction between Rna14p and Rna15p, we determined the minimal region in each protein required for association and the structure of the heterodimer in solution. The C-terminal residues of Rna14p wrap within the hinge region of Rna15p and form a locked molecular embrace. The structure explains the interdependent requirement of these two proteins and the molecular basis of several temperature-sensitive mutations. Furthermore it provides insight into the architecture and biology of Rna14p and Rna15p within the CF IA processing factor.

RESULTS

Minimal Rna14p/Rna15p Heterodimer

We prepared a series of truncated proteins to determine the minimal regions in Rna14p and Rna15p required for formation of a stable complex (Figure 1A). Based on detection of direct interaction of the expressed proteins (see Figure S1 available online), we obtained functional but size-reduced peptides containing residues 127–232 of Rna15p and residues 626–677 of Rna14p. We have also noted that high expression of either protein is achieved in *Escherichia coli* only on formation of a functional heterodimer. For example, further removal of residues 127–138 reduced heterodimer stability and consequently resulted in significantly lower expression levels (Figure S1). A similar interdependence of Rna14p and Rna15p stability was detected during overexpression studies in yeast (Bonneaud et al., 1994) and a similar codependence for CstF-77 and CstF-64 quantity was found in mammalian cells (Ruepp et al., 2011).

Coexpression Is Required to Form a Stable Folded Complex

Excellent stability and evidence of a single well-folded complex by nuclear magnetic resonance (NMR) and circular dichroism (CD) spectroscopy is seen with samples from bacterially coexpressed Rna14p$^{(626–677)}$/Rna15p$^{(127–232)}$ in which only Rna15p$^{(127—232)}$ initially harbors the N-terminal His_6-tag (Figures 1B and 1C). In contrast, expression and purification of isolated His_6-tagged Rna14p$^{(626–677)}$ or Rna15p$^{(127–232)}$ in *E. coli* produced only a limited amount of soluble protein. By CD spectroscopy, each of these isolated peptides is minimally folded (Figure 1B), with only minor indication of a thermal denaturation unfolding transition for His_6-Rna15p$^{(127–232)}$. NMR spectroscopy data are also consistent with peptides that are either unfolded or in a heterogeneous molten-globule state (Figure 1C). This is evident by the poor dispersion in the ^1H chemical shift values as is seen with disordered peptides, coupled with the inability to observe the majority of the observable backbone amide resonances that can be due to line-broadening effects of dynamics and multiple conformations. Only partial improvement was

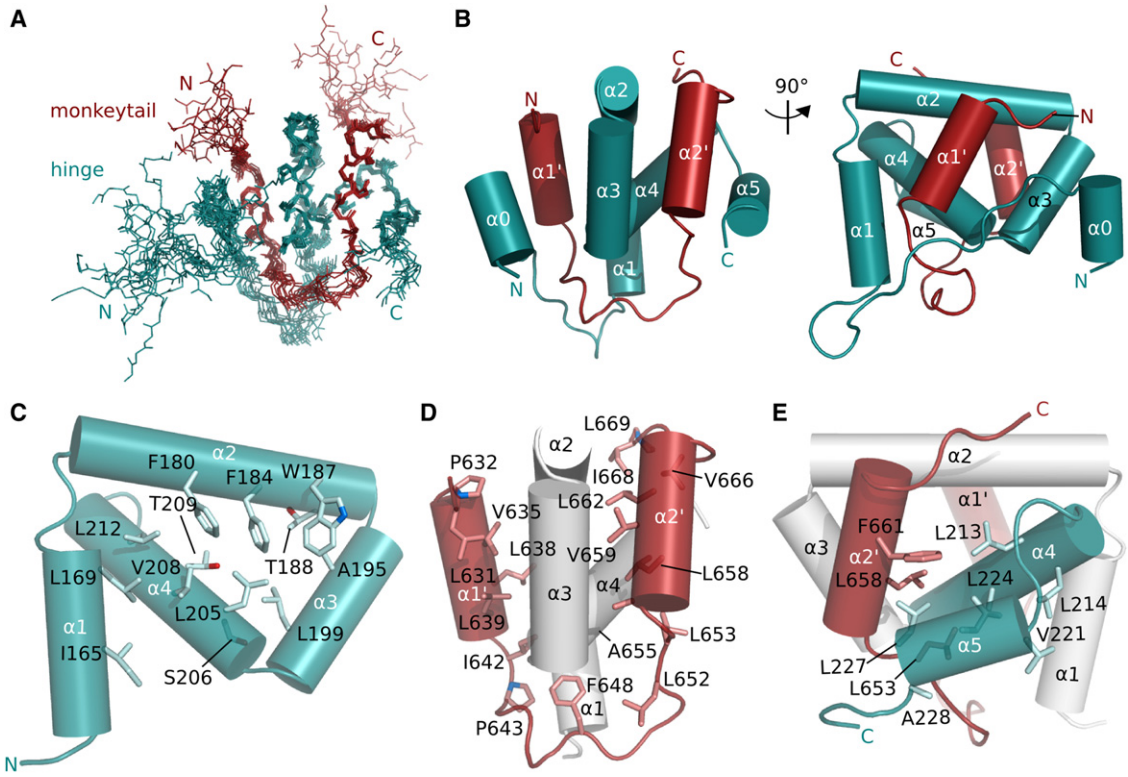

Figure 2. Structure of the Rna14p/Rna15p Heterodimer

(A) Ensemble of 10 lowest energy structures for Rna14p$^{(626–677)}$/Rna15p$^{(127–232)}$. The backbone atoms for each structure are depicted as sticks and visualized using PyMOL version 0.99 (PyMOL Molecular Graphics System, Schrödinger, LLC). The monkeytail domain from Rna14p is colored in dark red and the Rna15p hinge domain is colored in teal.

(B) Cartoon representation of the complex using the same coloring as in (A). At left, the orientation corresponds to that in (A), with the view at the right resulting from a 90° rotation from the left to right around the vertical axis. Helices are shown as cylinders and are annotated. A prime symbol is used to indicate helices from Rna14p.

(C) Intradomain side-chain contacts within the central core of Rna15p$^{(127–232)}$, residues Thr159 to Lys220. See Figure S2 for details of the buried polar residues.

(D) Residues in the monkeytail region of Rna14p$^{(626–677)}$ that make contact to Rna15p$^{(127–232)}$.

(E) Side-chain contacts from Rna15p helix α5 to helix α4 as well as to Rna14p helix α2′.

seen after the straightforward addition of equimolar quantities of His$_6$-Rna14p$^{(626–677)}$ and His$_6$-Rna15p$^{(127–232)}$. Moreover, the stability of the formed complex is poor and significant evidence of unfolded peptide remains (Figures 1B and 1C) and thus this method was not subsequently used to create Rna14p/Rna15p complexes.

Overall Description of the Rna14p/Rna15p Heterodimer Structure

Using a combination of distance, dihedral, and orientational restraints, we obtained a well-defined ensemble of structures for Rna14p$^{(626–677)}$/Rna15p$^{(127–232)}$ (Figure 2A and Table 1). The complex is formed around a central core of four α helices from Rna15p$^{(127–232)}$ (α1 to α4) that is encircled by the peptide from Rna14p$^{(626–677)}$ (Figure 2B). Additional N- and C-terminal helices from Rna15p complete the molecular embrace, including an extended clasp region located between helices α0 and α1 in Rna15p$^{(127–232)}$ that is devoid of regular secondary structure but which effectively locks and buries helix α1′ from Rna14p$^{(626–677)}$. The striking extent of interdomain contacts, and a resulting burial of 4900 ± 200 Å^2 solvent accessible surface (NACCESS; Hubbard and Thornton, 1993), at least partly explains the observed inability of the isolated Rna14p$^{(626–677)}$ or Rna15p$^{(127–232)}$ peptides to form a stable folded domain. Specific aspects of this bimolecular tether are detailed below.

Intradomain Contacts in the Central Core Region

An analysis of side-chain contacts reveals that those that occur between residues of the same peptide are essentially restricted to a central region within Rna15p$^{(127–232)}$ (Figure 2C). This hydrophobic volume is stabilized by aromatic residues along helix α2 (Phe180, Phe184, Trp187) as well as by aliphatic side chains from helices α1, α3, and α4 (Ile165, Leu169, Ala195, Leu199, Leu205, Val208, and Leu212). In addition, this region also contains the buried polar residues Thr188, Ser206, and Thr209. Because these residues are all contained within Rna15p$^{(127–232)}$, it is reasonable to predict that this central core could exist to some extent in the isolated peptide, and may explain the small thermal unfolding transition observed by CD spectroscopy (Figure 1B). However, the line-broadening and inability to observe most of the residues in the His$_6$-Rna15p$^{(127–232)}$ ^{15}N-HSQC (Figure 1C) indicates that this

Table 1. NMR and refinement statistics for the Rna14p/Rna15p minimal heterodimer

NMR distance and dihedral constraints	
Distance constraints	
Intraresidue	1226
Interresidue	
Sequential (|i − j| = 1)	604
Medium range (1 < |i − j| < 5)	454
Long range (|i − j| > 4)	229
Intermolecular	392
Ambiguous	1263
Hydrogen bonds	100 (50 × 2)
Total dihedral angle constraints	
Φ	114
Ψ	112
χ	42
Residual dipolar couplings	
$^1D_{HN}$	61
Structure statistics	
Violations (mean and SD)	
Distance constraints (Å)[a]	0.020 ± 0.001
Dihedral angle constraints (°)[b]	0.52 ± 0.07
Residual dipolar couplings, Q[c]	0.076 ± 0.006
Deviations from idealized geometry	
Bond lengths (Å)	0.004 ± 0.000
Bond angles (°)	0.59 ± 0.02
Impropers (°)	1.7 ± 0.1
Ramachandran plot, residues 138–228, 630–669 (%)[d]	
Residues in most favored regions	87.6
Residues in additionally favored regions	10.4
Residues in generally allowed regions	1.1
Residues in disallowed regions	0.9
Average pairwise rmsd, residues 138–228, 630–669 (Å)	
Heavy	0.66 ± 0.03
Backbone	0.37 ± 0.04

NMR, nuclear magnetic resonance; rmsd, root-mean-square deviation; SD, standard deviation.
[a] No violation >0.5 Å.
[b] No violation >5°.
[c] Calculated for each structure in the ensemble using the method described by Cornilescu et al. (1998).
[d] Determined by using PROCHECK-NMR (Laskowski et al., 1996), with residues in generally allowed or disallowed regions mainly restricted to the Rna14p loop region from Asn649 to Leu652.

structure, if formed, would be only partially stable or display conformational heterogeneity.

Polar Residues Buried within the Core of the Complex

A notable aspect of the heterodimer central region is the presence of buried polar residues in the otherwise hydrophobic core (Figure 2C). Residues such as Thr188, Ser206, and Thr209 are located away from the bulk aqueous solvent in the structure, and NMR spectroscopy confirms that these residue amides are distant from the accessible surface because they are unaffected by paramagnetic gadolinium in the sample buffer (Figure S2A) (Madl et al., 2009; Respondek et al., 2007). In addition, these side-chain hydroxyl protons do not exchange as usual with the solvent and are observable by NMR spectroscopy (Figure S2B). In several Rna15p sequences, and all metazoan hinge domains, Thr188, Ser206, and Thr209 are present as hydrophobic residues (Figure 3A). Mutation of the most central Ser206 and Thr209 to hydrophobic alanine remains compatible with complex formation via coexpression, with only a modest decrease on thermal stability for Thr209Ala, and a slight increase in stability for the Ser206Ala mutant (Figure S2C). The variable presence of these buried polar residues is therefore unclear. When present, the penalty due to loss of the hydrophobic side-chain appears to be compensated by an added contribution of the polar group to the hydrogen bonding network within the Rna14p/Rna15p heterodimer (Figure S2D). Scaffolding polar residues such as these are usually highly conserved (Worth and Blundell, 2009), such as Gln204 in the near absolutely conserved Pro-Gln-Leu motif that makes hydrogen bonds to backbone carbonyls within the complex (Figure 3A; Figure S2C). In contrast, the variable buried polar residues may serve a different role, possibly related to an aspect of coupled binding and folding, or to maintain protein solubility of any isolated and unfolded Rna15p.

Identification of the Rna14p Monkeytail Domain

Wrapping around the central core, the Rna14p$^{(626–677)}$ peptide makes extensive interdomain hydrophobic contacts to Rna15p (Figures 2B and 2D). We have named this previously uncharacterized Rna14p structural motif the monkeytail domain, as its mode of binding is reminiscent of a monkey tail grasping a tree branch. The C-terminal region of Rna14p that includes the monkeytail domain was previously shown to have sequence homology within the metazoan CstF-77 protein (Legrand et al., 2007; Simonelig et al., 1996). Using a combination of secondary structure prediction and the pattern of hydrophobic residues it is possible to clearly identify the corresponding region in orthologous yeast, fungi and metazoan Rna14/CstF-77 proteins (Figure 3B; Figure S3C). Within the yeast sequences, residues with interdomain contacts display 85% similarity and 40% identity, whereas the remaining residues are only 41% similar and 19% identical. Comparison of the metazoan sequences to that of *S. cerevisiae* reveal comparable levels of 87% similarity and 29% identity for the residues involved in contact with the hinge domain, and a striking low level of conservation for the remaining residues (17% similarity and zero identity). Although present at various distances from the HAT repeats and the C-terminal residue (Figure S3D), the monkeytail domain is nevertheless a conserved feature of the Rna14/CstF-77 proteins.

N and C Terminus of Rna15p$^{(127–232)}$ Complete the Molecular Embrace

The C-terminal helix α5 of Rna15p$^{(127–232)}$ makes intradomain contacts to helix α4 and through interdomain contacts partially covers helix α2′ of Rna14p$^{(626–677)}$ (Figure 2E). At the opposite side of the complex, a conformationally extended region from

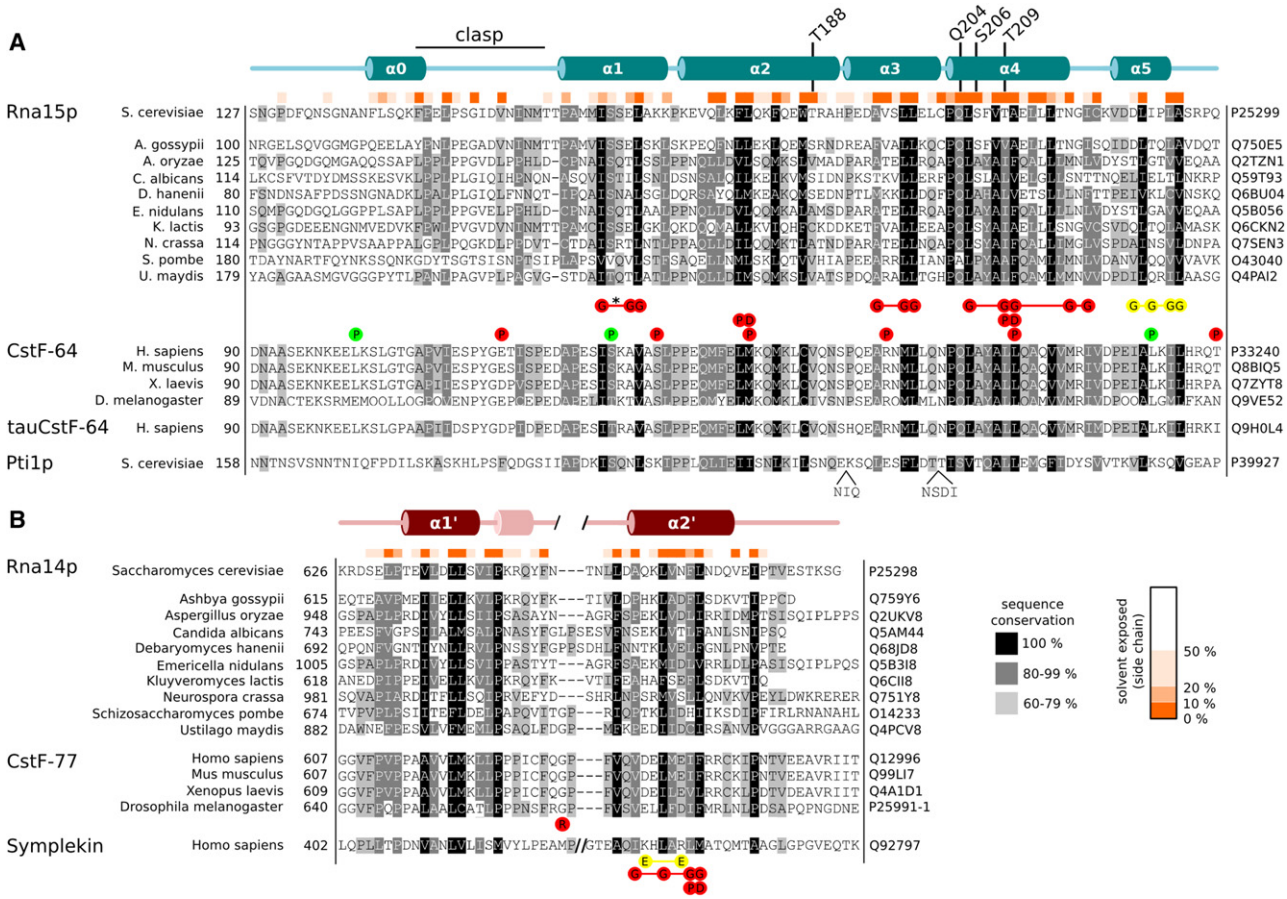

Figure 3. Sequence Conservation in the Monkeytail and Hinge Domains

(A) Sequence alignment of the hinge domain from various yeast and metazoan proteins. Along the top is a schematic representation of the secondary structure of Rna15p$^{(127–232)}$ within the *S. cerevisiae* heterodimer. The first residue in the alignment is numbered along the left, and the source of the sequence file in the database indicated along the right. Sequence conservation is grouped into 100% (black), 80%–99% (dark gray), and 60%–79% (light gray) based on the following similarity groups: acidic (DE), basic (KR), hydrophilic (HNQSTY), and hydrophobic (ACFILMPVW). The solvent accessibility of the side chain atoms is based on the *S. cerevisiae* Rna14p$^{(626–677)}$/Rna15p$^{(626–677)}$ structure calculated using NACCESS (Hubbard and Thornton, 1993), and colored dark orange (0%–10% accessible), orange (10%–20%), and light orange (20%–50%). Circles indicate known site-specific mutations from Hockert et al. (2010) and Ruepp et al. (2011), colored as no (green), moderate (yellow) and severe (red) perturbation. See Figure S3 for domain composition of the full-length proteins and secondary structure prediction of the hinge regions.

(B) Sequence conservation of the monkeytail domain, colored as in (A), with domains and secondary structure prediction in Figure S3. Mutations from Simonelig et al. (1996) and Ruepp et al. (2011).

residues Phe145 to Thr160 in Rna15p packs against helix α1′ of Rna14p$^{(626–677)}$, essentially locking the monkeytail peptide within the hinge domain helices via hydrophobic contacts and charge complementarity (Figure 4A). These regions, and indeed the entire section from helix α0 to the end of helix α5 in Rna15p$^{(127–232)}$ (residues 143–228), and residues Glu630 to Pro669 in Rna14p$^{(626–677)}$, display a relatively rigid backbone ({^{1}H}^{15}N heteronuclear NOE values >0.6) and T2 ^{15}N relaxation values consistent with a heterodimer of ~18 kDa (Figure 4B). This stable and locked arrangement is significant because it would prevent facile dissociation of the monkeytail peptide and implies that the complex, once formed, is incapable of simple separation. Such a model has important implications in the biology of the Rna14p and Rna15p proteins because the heterodimer would be inaccessible to dissociation or potential rebinding to additional partners in a stepwise or otherwise regulated manner. As a result we chose to investigate the stability of this clasp region with respect to formation of the bimolecular complex.

The Clasp Region Can Locally Unfold

CD spectroscopy indicates a global unfolding midpoint transition at 3.2 M urea (Figure 4C). Using NMR spectroscopy, we analyzed the same urea-induced unfolding of the complex at the residue level, and observed any conformational denaturing by the movement of backbone amide cross peaks as a function of added urea (Figure 4D; Figure S4). At urea concentrations up to 2M only residues from Ser142 to Ser150 are significantly affected by the denaturant, and thus the data indicate an initial local unfolding of the central clasp region preceding the global unfolding of the complex. It is therefore possible that to a small degree under suitable conditions, posttranslational modification or in the presence

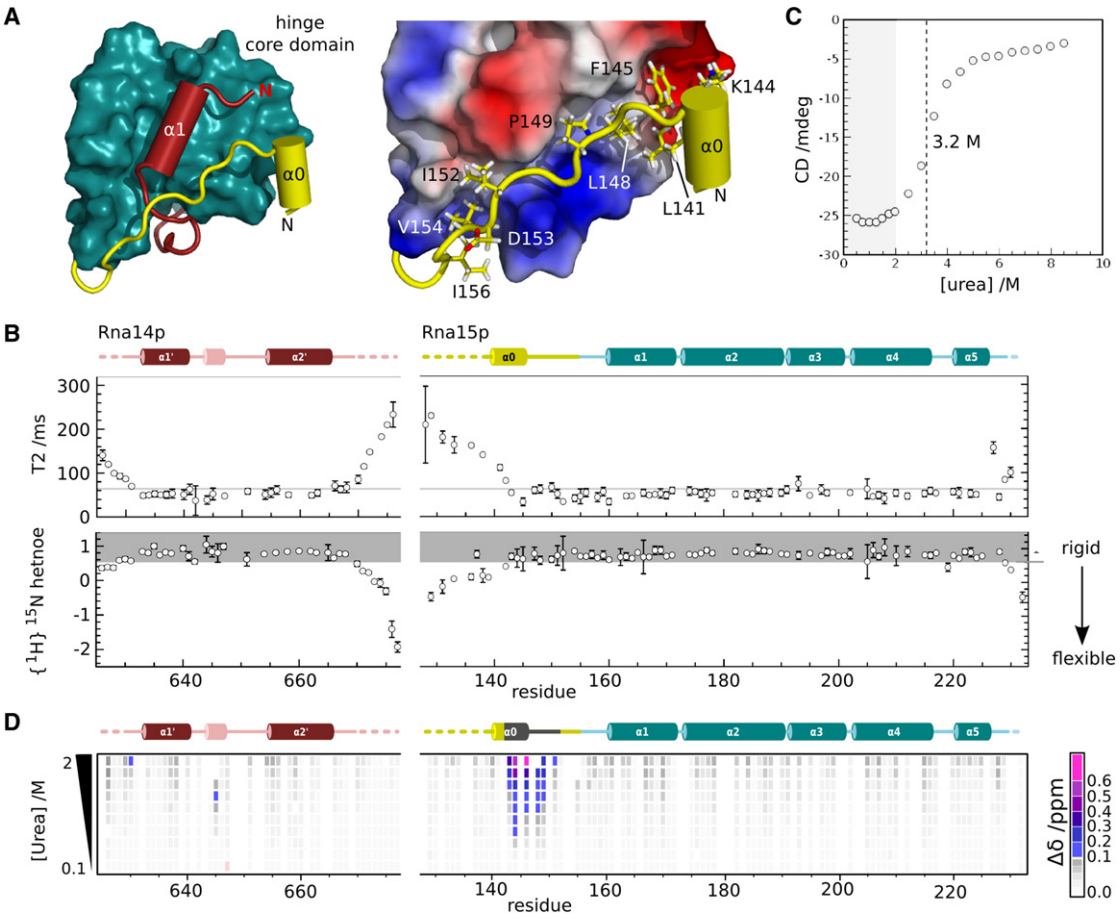

Figure 4. Structural Investigation of the Clasp Region

(A) Left, cartoon representation of the clasp region (yellow) crossing over the Rna14p$^{(626–677)}$ peptide (dark red) and against the hinge core domain (teal surface). Right, clasp region contacting residues involve hydrophobic interactions as well as charge complementarity. The surface is colored by using PyMOL and a scale from negative (red) to positive (blue).

(B) Backbone conformational flexibility of [^{15}N]-Rna14p$^{(626–677)}$/Rna15p$^{(127–232)}$ measured by NMR spectroscopy ^{15}N relaxation at 303 K and 700 MHz. Transverse (T2) relaxation and {^{1}H}^{15}N heteronuclear NOE values are depicted as gray-filled circles and error bars representing the mean ± standard error of mean for three independent measurements. For T2 measurements, lower values are typically indicative of larger and less conformationally flexible backbone amides. The horizontal line corresponds to the T2 value (65 ms) predicted for a folded globular protein of the same molecular weight as the heterodimer (Daragan and Mayo, 1997). For the {^{1}H}^{15}N -heteronuclear NOE (hetnoe) measurements, values >0.6 (within the gray shaded region) indicate reduced conformational flexibility.

(C) CD spectroscopy at 222 nm was used to monitor unfolding of 15 μM Rna14p$^{(625–677)}$/Rna15p$^{(127–232)}$ due to increasing urea concentration, at a temperature of 303 K and a buffer of 20 mM sodium phosphate (pH 6.5) and 50 mM NaCl. The midpoint unfolding transition is observed at 3.2 M urea.

(D) Urea-induced local unfolding of the region around helix α0 in Rna15p. Increasing amounts of urea were added to a 200 μM sample of [^{15}N]-Rna14p$^{(626–677)}$/Rna15p$^{(127–232)}$ and a ^{1}H^{15}N-HSQC collected at each point at 303 K and 700 MHz (see Figure S4). The change in chemical shift at each point, $\Delta\delta = ((\Delta\delta_{15N})^2 + (5 \times \Delta\delta_{1H})^2)^{0.5}$, was calculated using the initial HSQC as a control spectrum. The graph includes rectangles for each residue (along the x axis) at each concentration from 0.1 to 2 M urea (y axis), colored by the amount of chemical shift perturbation. Only the region from 142 to 150 (blue to magenta-colored rectangles) is affected by urea concentrations up to 2 M.

of binding partners, that the clasp region could be displaced and allow for dissociation of Rna14p and Rna15p. Sequence similarity in the clasp regions from orthologous Rna15p (Figure 3B) especially for the buried hydrophobic residues corresponding to Rna15p Phe145, Leu148, Pro149, Ile152, Val154, Ile156 (Figure 4A, right) suggest that this locking element from residue 145 to 158 is a common feature of the yeast/fungi monkeytail-hinge domain tether. The presence of this clasp in metazoa is not as clear, although there is conservation of hydrophobic residues corresponding to Phe145, Leu148, and Ile156, and a mutation in this region for human CstF-64 impairs complex formation (Glu117Pro in Figure 3B) (Hockert et al., 2010). In contrast to the clasp region, there is only a low degree of conservation within the N-terminal residues up to and including helix α0, and this region in Rna15p within the Rna14p$^{(626–677)}$/Rna15p$^{(127–232)}$ complex displays notable backbone conformational flexibility (Figures 2A and 4B). Secondary structure predictions also fail to indicate a consistent helical propensity in this region for other Rna15p and CstF-64 proteins (Figure S3A), implying that the helix α0 element may be variably present.

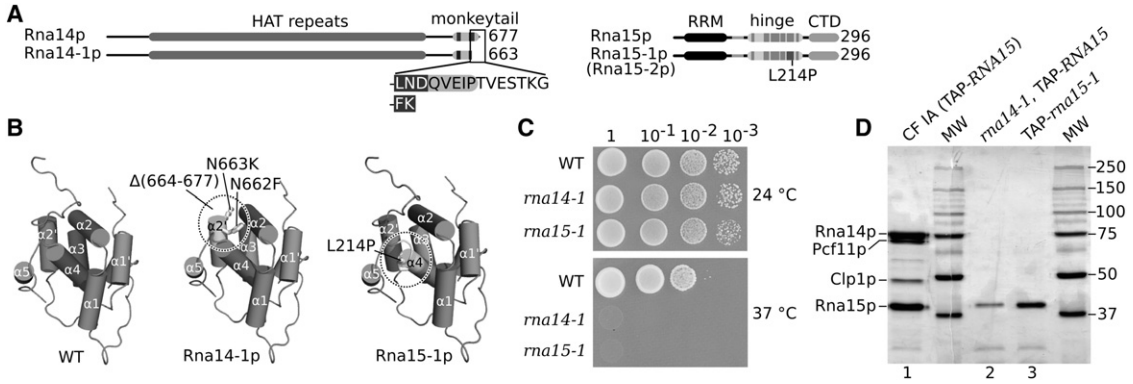

Figure 5. Stability of CF IA Composition in *rna14-1* and *rna15-1* Mutants

(A) Schematic representation of wild-type and mutant *S. cerevisiae* Rna14p and Rna15p. Rna14-1p includes a C-terminal truncation within the monkeytail domain (Rouillard et al., 2000). Rna15-1p (equivalent to Rna15-2p found in an independent screen) contains leucine to proline conversion at residue 214 (Qu et al., 2007).

(B) Location of mutations mapped onto the structures of Rna14p$^{(626–677)}$/Rna15p$^{(127–232)}$.

(C) TAP-tagged Rna15p-expressing strains were tested for temperature-sensitivity. Serial dilutions of wild-type (WT) and mutant *rna14-1* and *rna15-1* cells expressing TAP-tagged Rna15 proteins were spotted on rich medium and incubated at 24°C or 37°C. Temperature-sensitivity correlates with the presence of the mutant allele and is not due to the addition of the TAP-tag.

(D) Silver-stained 10% polyacrylamide SDS gel of protein factors prepared from TAP-tagged yeast strains. CF IA was purified from TAP-tagged Rna15p in either the wild-type (lane 1) or the temperature sensitive *rna14-1* (lane 2) or *rna15-1* (lane 3) mutant backgrounds. MW, molecular weight marker in kDa.

Incorporation of Rna15p into CF IA Requires a Structurally Stable Monkeytail/Hinge Association

Previous studies identified mutants in Rna14p and Rna15p that inhibit proper 3′ pre-mRNA processing and produce a temperature-sensitive phenotype. Interestingly, the early and commonly used alleles *rna14-1* and *rna15-1* have mutations that we now locate within the Rna14p monkeytail and Rna15p hinge regions (Figures 5A and 5B). In *rna14-1*, a frame-shift mutation converts the final 16 residues of the protein from -LNDQVEIPTVESTKG to only two residues, -FK (Rouillard et al., 2000). This mutation truncates part of helix α2′ within the monkeytail domain. The protein encoded by *rna15-1* (equivalent to *rna15-2*) has a disrupted helix α4 due to a single amino acid replacement at position 214 from leucine to proline, and likely perturbs contacts between helix α4 and α5 in Rna15p (Figure 2E) (Qu et al., 2007). Because the mutations from either *rna14-1* or *rna15-1* would specifically destabilize the minimal heterodimer, we predict that protein dissociation would also occur with the corresponding full-length mutant proteins. Using the TAP-tag approach (Puig et al., 2001, Rigaut et al., 1999) we generated strains containing either TAP-tagged wild-type Rna15p within the *rna14-1* mutant, or TAP-tagged Rna15p(L214P) in the *rna15-1* mutant. The constructed strains display the predicted growth defect at 37°C and normal growth at the permissive temperature of 24°C (Figure 5C). Purification using the TAP-tag protocol reveals normal CF IA composition using TAP-tagged Rna15p wild-type cells (Figure 5D, lane 1). In contrast, purification from cells grown at permissive temperature of TAP-Rna15p from *rna14-1* mutant (lane 2), or purification of TAP-Rna15p(L214P) in *rna15-1* mutant (lane 3) results in isolated Rna15 protein and no other CF IA subunits. Therefore it appears that in keeping with the structural information, the main consequence of *rna14-1* or *rna15-1* is disrupted tethering and dissociation of Rna15p from CF IA, although the association state of Rna14p with either Pcf11p or Clp1p is not known.

Architecture of Full-Length Rna14p and Rna15p

Genetic data confirm that mutation of residues within the minimal heterodimer region inhibit association of full-length Rna14p and Rna15p. However, the structural contribution of the heterodimer region to the larger Rna14p and Rna15p constructs were as yet unknown. We therefore prepared a series of complexes containing different sized peptides and investigated their properties using NMR spectroscopy.

The chemical shift or, equivalently, the amide crosspeak position of a specific residue in a NMR spectrum (such as the two-dimensional ^{1}H,^{15}N-HSQC) is exquisitely sensitive to the local environment of the residue. Changes in the amide peak position therefore reflect subtle differences in structure, presence of added ligands or changes in protein dynamics. Comparison of the spectra between the minimal heterodimer and that of the full-length Rna15p bound to Rna14p$^{(626–677)}$ shows significant similarity (Figure 6A, red and blue spectra, respectively), and thus suggests that the same monkeytail-hinge domain structure is also present within this larger context. The isolated N-terminal RRM domain of Rna15p (residues 1–104) (Figure 6A, green spectrum) also shows excellent similarity to the full-length Rna15p/Rna14p$^{(626–677)}$ spectra (blue), revealing that the structure of the RRM domain remains the same within the full-length Rna15p and is unchanged on the association with the Rna14p monkeytail domain.

The NMR spectrum for the complete Rna14p/Rna15p complex is shown in Figure 6B (black spectrum), and is surprising because although the complex is prohibitively large for standard NMR investigation (the predicted tetramer assembly is 226 kDa) there are numerous observable cross peaks in the standard ^{15}N-HSQC spectrum. The majority of these peaks show exact correlation to that of the Rna15p RRM domain (green spectrum). Due to the overall size of the complex, this result indicates that with the full-length Rna14p/Rna15p

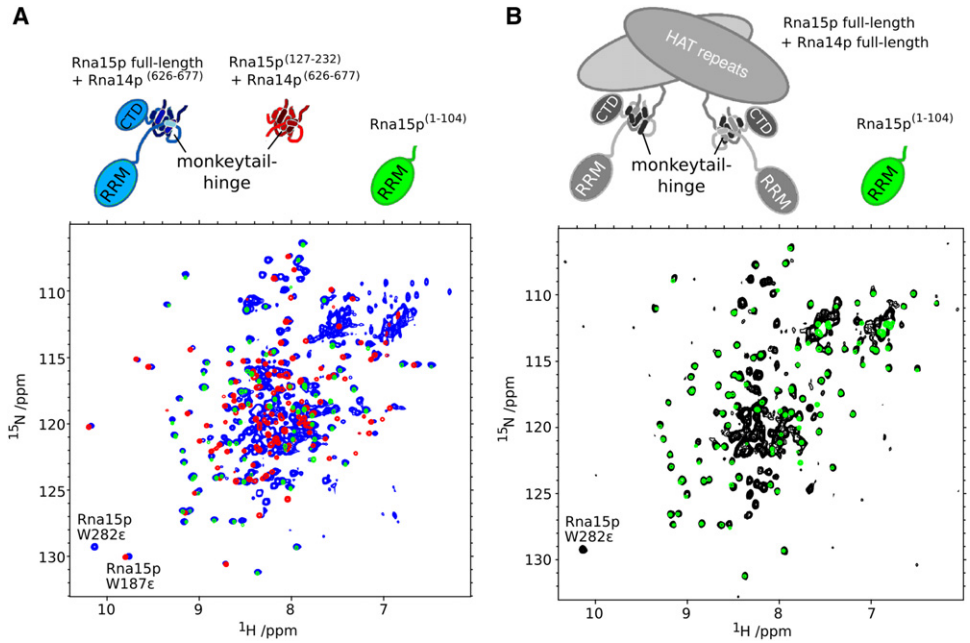

Figure 6. Architecture of Larger Rna14p and Rna15p Constructs

(A) Overlay of $^{1}H,^{15}N$-TROSY spectra of [^{15}N]-Rna14p$^{(626–677)}$/Rna15p (full-length) (blue), minimal heterodimer ([^{15}N]-Rna14p$^{(626–677)}$/Rna15p$^{(127–232)}$; red) and the RRM from Rna15p ([^{15}N]-Rna15p$^{(1–104)}$; green). The two tryptophan $H^{\varepsilon 1}N^{\varepsilon 1}$ cross peaks are annotated by residue number.

(B) Overlay of $^{1}H,^{15}N$-HSQC spectra of full-length [^{15}N]-Rna14p/Rna15p (black), and the RRM from Rna15p ([^{15}N]-Rna15p$^{(1–104)}$; green). The tryptophan $H^{\varepsilon 1}N^{\varepsilon 1}$ crosspeak that may belong to the C-terminal domain of Rna15p (Trp282) is indicated.

complex, the structure of the RRM domain is the same as in isolation and also demonstrates significant independent mobility, possibly enabled through a connection to the larger core particle of the complex solely through a flexible linker (such as via the opa region) (Figure S3B). In contrast, all of the cross peaks relating to the hinge domain and monkeytail are not observed in the spectrum of the full-length complex (Figure 6B). The peaks are broadened beyond detection, which is consistent with incorporation of the monkeytail-hinge tether into the large-sized core structure. The remaining peaks observed in the full-length Rna14p/Rna15p tetramer (black spectrum) likely arise from the C-terminal domain of Rna15p that may also display independent mobility within the large tetramer.

DISCUSSION

The interaction between the monkeytail and hinge domains creates a high affinity link that joins Rna14p and Rna15p together in a tight complex, and from sequence analysis a similar tether is likely formed between the metazoan orthologs CstF-77 and CstF-64. The solution structure of the minimal heterodimer reveals the molecular basis for the association. This intimate embrace of the two peptides differs from a typical interaction between two protein domains, in which the association areas are usually restricted to small defined surface regions. It also differs from the usual coupling between binding and folding in which only one of the components is natively unstructured (Wright and Dyson, 2009). As a consequence, the final stabilizing hydrophobic core is jointly contributed by each of the two domains, and isolated monkeytail or hinge domains are unable to form a stable fold by themselves. The three-dimensional tether structure is likely unchanged within the full-length protein complex, where it appears to be in relatively close contact with the main core of the tetrameric Rna14p/Rna15p. In contrast, the Rna15p RRM domains, and possibly the Rna15p C-terminal regions, retain motional independence in all constructs. A model of the association between Rna14p and Rna15p consistent with the data is presented in Figure 7.

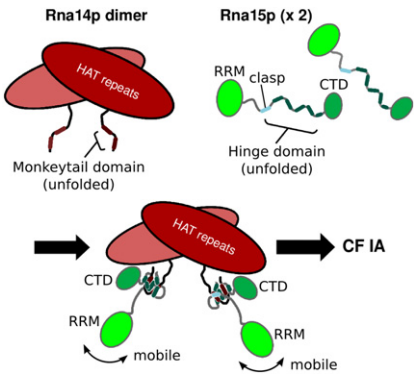

Figure 7. Model of the Monkeytail and Hinge Domains in the CF IA Complex

Schematic representation of the association between the Rna14p dimer and two molecules of Rna15p in the creation of the CF IA complex. The RRM domain, and possibly the carboxy-terminal domain (CTD), exhibits independent mobility within the complex.

Structural Basis of Processing Defects in *rna14* and *rna15* Mutants

The atomic details of the minimal heterodimer indicate that both *rna14-1* and *rna15-1* mutants act to destabilize the association between Rna14p and Rna15p. Using TAP-tag purification we have found that the loss of this tether specifically results in a separation of Rna15p from CF IA. The cause of defects in *rna14-1* and *rna15-1* can therefore be re-evaluated in light of these findings. Most importantly, without Rna15p the RRM domain is no longer present in the CF IA complex, causing loss of specific RNA-binding for CF IA, even though the RRM domain itself is not directly affected by either mutation.

In isolation, Rna15p does not bind RNA with significant specificity to the pre-mRNA cleavage site (Gross and Moore, 2001a) and wild-type Rna15p does not crosslink to RNA in the *rna14-1* mutant (Kim et al., 2004). Functional interaction with Rna14p and incorporation into CF IA, or addition of Hrp1p, is required for increased RNA-binding affinity and specificity (Noble et al., 2004; Leeper et al., 2010). This may arise for a number of reasons. First, because Rna14p forms a homodimer and each Rna14p is tightly bound to Rna15p, this gives a stoichiometry of 2:2 Rna14p/Rna15p in wild-type CF IA. This arrangement could lead to synergistic binding by the two Rna15p RRM domains to tandem nucleotide sequences in the pre-mRNA that might enhance both the affinity and the specificity of the interaction. Second, contacts between the Hrp1p and Rna15p RRM domains create specificity for adjacent enhancer and positioning element sequences, respectively, within the composite anchor element (Leeper et al., 2010) and Rna14p provides an important bridging contact to both Rna15p and Hrp1p to stabilize this ternary complex (Gross and Moore, 2001a). The result in either case is destabilization of the Rna14p-Rna15p interaction in the *rna14-1* and *rna15-1* mutants at nonpermissive temperatures, loss of Rna15p from CF IA, and thus the reduced RNA-binding fails to target a specific site for cleavage. The lack of stable downstream processing complexes leads to failed termination of RNA polymerase II activity, thus producing extended pre-mRNA transcripts, perhaps until collision with a polymerase traveling in the opposite direction (Torchet et al., 2002).

The fidelity of the Rna14p-Rna15p interaction is also important for stability of the two proteins in the yeast cell. Overexpressed Rna14p or Rna15p can only accumulate in yeast when both are overexpressed (Hammell et al., 2002). Rna15p overexpression can serve to restore normal growth to *rna14-1* mutants (and vice versa), probably by partially overcoming the reduced K_D with increased protein concentration (Minvielle-Sebastia et al., 1991). Similarly, mutation of the STS1/DBF8 gene, a factor in the ubiquitin/proteasome pathway, seems to help *rna15-1* and not *rna14-1* by increasing Rna15p levels through reduction in the usual degradation mechanism (Amrani et al., 1996).

It is also possible that the mutations prevent or promote functional interaction with additional binding partners. Opposing this hypothesis, all three lanes in Figure 5D display no evidence of alternate stable complexes formed by the TAP-tagged Rna15p apart from the CF IA subunits in the wild-type strain. However, low affinity or transient interactions cannot be ruled out, nor can the presence of additional complexes formed with Rna14p. The Pti1p protein, for example, has a similar composition to Rna15p (Figure 3A; Figures S3A and S3B) and via a region similar to the hinge domain interacts with Rna14p (Nedea et al., 2003, Skaar and Greenleaf, 2002).

Implications for the Metazoan CstF

Similarity of sequence and predicted secondary structure for the monkeytail regions of Rna14p and CstF-77 (Figure 3B; Figure S3C), as well as the hinge domains from Rna15p and CstF-64 (Figure 3A; Figure S3A), suggests a similar tether structure will form between CstF-77 and CstF-64. Accordingly, the hinge domain of CstF-64 was previously characterized as a conserved region necessary for interaction with CstF-77 (Bai et al., 2007; Hockert et al., 2010) and this region is also conserved in the paralogue τCstF-64 (Figure 3A). A significant number of hinge mutations within human CstF-64 result in loss of protein function, although it appears that helix α5 and to some extent helix α1 may demonstrate less importance in the formation of the tether structure (Figure 3A) (Hockert et al., 2010, Ruepp et al., 2011). In addition, the *Drosophila* suppressor of forked (dCstF-77) R-9-18 mutant harbors a Gly664Arg exchange located in the loop between helix α1′ and α2′ of the monkeytail region (Simonelig et al., 1996). The mutant would likely destabilize but not necessarily prevent tether formation, thus leading to the observed phenotype.

In contrast to the Rna14p and Rna15p sequences, the metazoan orthologs display a remarkable level of conservation throughout the monkeytail and hinge domains, and thus present highly retained surface residues. This finding could indicate that, unlike the yeast proteins, the surface of the monkeytail-hinge tether could also be important in binding components of the CstF subunits or other factors. These secondary contacts would only be possible after formation of the folded tether complex, and thus may provide an interesting mechanism to prevent binding to isolated CstF-64 or CstF-77.

Presence of Monkeytail/Hinge Domain-Like Interactions in Other Systems

A significant finding from the structure of the Rna14p/Rna15p complex is the remarkable level of surface contact between the two proteins, and the accompanying link between binding and folding. Given the identification of this unique association mechanism, it is of interest to see if a similar interaction mode is present in other protein complexes. Unfortunately a general sequence search for additional monkeytail or hinge domains was not successful due to several factors. Even among the orthologous proteins listed in Figure 3, the sequence conservation of the monkeytail and hinge domains are low. To further complicate the analysis, in several cases polar residues occur within the otherwise conserved hydrophobic residues (such as the buried Thr188, Ser206, and Thr209 in Rna15p), although this trend appears to be more pronounced within the yeast proteins than among the metazoan sequences. Given the tight and specific association, as well as the large binding surface, it is surprising that the component peptides display such a high level of sequence variation. However, it is equally apparent that valuable information can be extracted from the striking pattern of hydrophobic residues due to the buried residues with the complex. We have used this approach to identify regions in yeast Pti1p and human symplekin that may adopt similar structures to the

Rna15p hinge and Rna14p monkeytail peptides (Figure 3; Figure S3).

Future attempts to identify similar monkeytail-hinge tethers in other systems may successfully combine secondary structure prediction, sequence alignment and high-resolution data of protein segment interactions (including domain-based interactome studies, such as Boxem et al. [2008]). The necessity for domains such as the hinge and monkeytail to bind to form a stable fold also has implications for automated structure-prediction algorithms, such as threading and other modeling protocols. Identification of other systems will also lead to better understanding of the parameters governing dipeptide-dependent folding and dimerization, and the creation of specific, strong, and stable protein tethers.

EXPERIMENTAL PROCEDURES

Protein Expression and Purification

The DNA encoding yeast Rna14p or Rna15p was cloned as several construct lengths by using *NdeI* and BamHI restriction sites introduced by PCR, into pET-derived plasmids as either untagged or N-terminal His_6-tagged constructs, with either streptomycin or ampicillin resistance, respectively (Romier et al., 2006). Coexpression was necessary to produce all heterodimeric Rna14p/Rna15p complexes, using two vectors and media containing both streptomycin and ampicillin. Expression and purification were performed as described (Mackereth, 2011). In brief, plasmids were used to transform BL21(DE3) or BL21 *lysY* (New England BioLabs) with growth in LB or M9 minimal media supplemented with 1 g × L^{-1} $^{15}NH_4Cl$ and/or 2 g × L^{-1} [^{13}C] glucose. Leucine-specific ^{13}C-labeling used 100 mg × L^{-1} [^{13}C,^{15}N]leucine, and methyl-specific protonation of an otherwise uniformly-deuterated sample was prepared as described (Goto et al., 1999). Purification utilized Ni^{2+}-affinity chromatography in a buffer of 50 mM Tris, pH 7.5, 500 mM NaCl, 5% (v/v) glycerol with 5 mM, 25 mM, or 250mM imidazole added for the binding, wash and elution steps, respectively. The His_6-tag was removed by adding Tobacco Etch Virus to the samples in phosphate buffered saline (PBS) and after a final passage through the Ni^{2+}-affinity column, the purified protein was concentrated and exchanged to 20 mM sodium phosphate, pH 6.5, 50 mM NaCl, and 2 mM dithiothreitol.

NMR Spectroscopy

Samples contained between 0.2 to 1 mM protein in 20 mM sodium phosphate, pH 6.5, 50 mM NaCl, 2 mM dithiothreitol, and either 10% or 99% ^2H_2O. NMR experiments were conducted at a temperature of 303 K on a triple resonance Bruker Avance 700 MHz, as well as a cryoprobe-equipped Bruker Avance 800 MHz and Varian 800 MHz. Spectra were processed using NMRPipe/Draw (Delaglio et al., 1995) and analyzed using Sparky 3 (T.D. Goddard and D.G. Kneller, University of California, San Francisco, CA). Chemical shift assignments were completed as described (Mackereth, 2011). Proton distances were obtained from a 3D ^1H,^{15}N-HSQC-NOESY (100 ms mixing time), and from ^1H,^{13}C-HSQC-NOESY spectra (120 ms mixing time) using either uniformly ^{13}C- or leucine ^{13}C-labeled protein, or methyl protonated Leu, Val, Ile(δ1) in an otherwise ^2H,^{13}C,^{15}N-labeled protein. ^1H-^{15}N residual dipolar couplings were measured using an interleaved spin state-selective ^1H,^{15}N-TROSY experiment, without and with ~18 mg/ml Pf1 phage (Profos AG, Regensberg, Germany). Amide ^{15}N relaxation data were acquired at 600 MHz and 303 K as described (Farrow et al., 1994). Steady-state heteronuclear {^1H}^{15}N-NOE spectra were recorded with and without 3 s of ^1H saturation. Relaxation rates and error calculations were determined using NMRView v. 4 (Johnson and Blevins, 1994).

Structure Calculation

The ensemble of structures was calculated by using a standard ARIA 1.2/CNS 1.1 setup (Brünger et al., 1998, Nilges, 1995). A set of 2814 manually-assigned distance restraints was accompanied by peaks with ambiguous assignment from the four NOESY spectra. Hydrogen bond restraints (two per hydrogen bond) were included for amides that displayed reduced exchange with ^2H_2O and were within a clearly identified α-helix. Dihedral angles were obtained from TALOS+ (Shen et al., 2009), and used only if the values were consistent with those predicted from DANGLE (Cheung et al., 2010). The error was set at twice the TALOS error, but with a minimum error set at 20°. Leucine χ angles were only used when $^{13}C^{\delta1}$ and $^{13}C^{\delta2}$ values were indicative of a single rotamer (Mulder, 2009). Residual dipolar couplings for residues with {^1H}^{15}N heteronuclear NOE values >0.6, were incorporated starting at iteration four, and used Da and R values of 11.0 and 0.5, respectively. The 10 lowest energy structures after the final water refinement of 20 structures were taken as the calculated ensemble.

Solvent paramagnetic relaxation enhancement

Amide proton T1 values were measured via saturation-recovery 2D ^1H,^{15}N-HSQC spectra (Madl et al., 2009) at 303 K using a Bruker Avance 700 MHz spectrometer with relaxation delays of 0.1, 0.2, 0.3, 0.5, 0.75, 1, 2, and 4 s. Purified gadolinium diethylenetriamine-pentaacetic acid (Respondek et al., 2007) was added at concentrations of 0, 1, 2, 3, 5, 7, and 10 mM to 500 uL of 0.8 mM protein in 20 mM sodium phosphate, pH 6.5, 50 mM NaCl and 10% (v/v) ^2H_2O. Relaxation rates were calculated using NMRViewJ (One Moon Scientific).

Circular Dichroism

Spectra were measured using a Jasco-815 CD spectrometer equipped with a PFD-425S/15 Peltier thermal control. Samples were measured at 15 μM in 20 mM sodium phosphate (pH 6.5), 50 mM NaCl and varying concentrations of urea, using a 1 mm quartz cell. Wavelength scans were collected as an average of three scans performed at 30°C with a scanning speed of 50 nm × min^{-1} and data acquisition for 0.5 s every 1 nm. Thermal denaturation studies at 222 nm used a heating rate of 1°C × min^{-1}, with data sampled for 4 s every 2°C with a bandwidth of 1 nm.

Yeast Experiments

CF IA was purified from isogenic yeast strains where the Rna15 protein was TAP-tagged at its N terminus according to Dheur et al. (2005). Wild-type CF IA was obtained from strain YSD12 (TAP::*RNA15*, *ade2 leu2 ura3 trp1Δ his3*). Mutant forms of the factor were produced from strains YLM196 (TAP::*RNA15, rna14-1 ade2 leu2 ura3 trp1Δ his3*) and YLM197 (TAP::*rna15-1*, *ade2 leu2 ura3 trp1Δ his3*).

ACCESSION NUMBERS

Coordinates have been deposited in the PDB (2L9B).

SUPPLEMENTAL INFORMATION

Supplemental Information includes four figures and can be found with this article online at doi:10.1016/j.str.2011.02.003.

ACKNOWLEDGMENTS

We thank Tobias Madl for assistance with measurement and analysis of solvent PRE data. Financial support from the TGE RMN THC Fr3050 for conducting the research is gratefully acknowledged. The eNMR project (European FP7 e-Infrastructure grant, contract no. 213010, www.enmr.eu), supported by the national GRID Initiatives of Italy, Germany and the Dutch BiG Grid project (Netherlands Organization for Scientific Research), is acknowledged for the use of web portals, computing, and storage facilities. The project was supported by funds from the Institut de Chimie des Substances Naturelles (C.D.M.), the Aquitaine regional government (S.F., C.D.M), La Fondation pour la Recherche Médicale/Fondation BNP-Paribas (to L.M.-S.), the Institut Européen de Chimie et Biologie (S.F., C.D.M.), by an Avenir/ARC grant from INSERM (S.F.) and PhD fellowships from the Caixa foundation, and the Association pour la Recherche sur le Cancer (M.M.M.).

Received: January 7, 2011
Revised: February 11, 2011
Accepted: February 13, 2011
Published: April 12, 2011

REFERENCES

Amrani, N., Dufour, M.E., Bonneaud, N., and Lacroute, F. (1996). Mutations in STS1 suppress the defect in 3′ mRNA processing caused by the rna15-2 mutation in Saccharomyces cerevisiae. Mol. Gen. Genet. *252*, 552–562.

Bai, Y., Auperin, T.C., Chou, C.-Y., Chang, G.-G., Manley, J.L., and Tong, L. (2007). Crystal structure of murine CstF-77: dimeric association and implications for polyadenylation of mRNA precursors. Mol. Cell *25*, 863–875.

Birse, C.E., Minvielle-Sebastia, L., Lee, B.A., Keller, W., and Proudfoot, N.J. (1998). Coupling termination of transcription to messenger RNA maturation in yeast. Science *280*, 298–301.

Bonneaud, N., Minvielle-Sebastia, L., Cullin, C., and Lacroute, F. (1994). Cellular localization of RNA14p and RNA15p, two yeast proteins involved in mRNA stability. J. Cell Sci. *107*, 913–921.

Boxem, M., Maliga, Z., Klitgord, N., Li, N., Lemmens, I., Mana, M., de Lichtervelde, L., Mul, J.D., van de Peut, D., Devos, M., et al. (2008). A protein domain-based interactome network for C. elegans early embryogenesis. Cell *134*, 534–545.

Brünger, A.T., Adams, P.D., Clore, G.M., DeLano, W.L., Gros, P., Grosse-Kunstleve, R.W., Jiang, J.S., Kuszewski, J., Nilges, M., Pannu, N.S., et al. (1998). Crystallography & NMR system: a new software suite for macromolecular structure determination. Acta Crystallogr. D Biol. Crystallogr. *54*, 905–921.

Cheung, M.S., Maguire, M.L., Stevens, T.J., and Broadhurst, R.W. (2010). DANGLE: a Bayesian inferential method for predicting protein backbone dihedral angles and secondary structure. J. Magn. Reson. *202*, 223–233.

Cornilescu, G., Marquardt, J.L., Ottiger, M., and Bax, A. (1998). Validation of protein structure from anisotropic carbonyl chemical shifts in a dilute liquid crystalline phase. J. Am. Chem. Soc. *120*, 6836–6837.

Danckwardt, S., Hentze, M.W., and Kulozik, A.E. (2008). 3′ end processing: molecular mechanisms and implications for health and disease. EMBO J. *27*, 482–498.

Daragan, V.A., and Mayo, K.H. (1997). Motional analyses of protein and peptide dynamics using ^{13}C and ^{15}N NMR relaxation. Prog. Nucl. Magn. Reson. Spectrosc. *32*, 63–105.

Delaglio, F., Grzesiek, S., Vuister, G.W., Zhu, G., Pfeifer, J., and Bax, A. (1995). NMRPipe: a multidimensional spectral processing system based on UNIX pipes. J. Biomol. NMR *6*, 277–293.

de Vries, H., Rüegsegger, U., Hübner, W., Friedlein, A., Langen, H., and Keller, W. (2000). Human pre-mRNA cleavage factor II(m) contains homologs of yeast proteins and bridges two other cleavage factors. EMBO J. *19*, 5895–5904.

Dheur, S., Nykamp, K.R., Viphakone, N., Swanson, M.S., and Minvielle-Sebastia, L. (2005). Yeast mRNA Poly(A) tail length control can be reconstituted in vitro in the absence os Pab1p-dependent Poly(A) nuclease activity. J. Biol. Chem. *280*, 24532–24538.

Farrow, N.A., Muhandiram, R., Singer, A.U., Pascal, S.M., Kay, C.M., Gish, G., Shoelson, S.E., Pawson, T., Forman-Kay, J.D., and Kay, L.E. (1994). Backbone dynamics of a free and phosphopeptide-complexed Src homology 2 domain studied by 15N NMR relaxation. Biochemistry *33*, 5984–6003.

Gavin, A.C., Bösche, M., Krause, R., Grandi, P., Marzioch, M., Bauer, A., Schultz, J., Rick, J.M., Michon, A.M., Cruciat, C.M., et al. (2002). Functional organization of the yeast proteome by systematic analysis of protein complexes. Nature *415*, 141–147.

Goto, N.K., Gardner, K.H., Mueller, G.A., Willis, R.C., and Kay, L.E. (1999). A robust and cost-effective method for the production of Val, Leu, Ile (delta 1) methyl-protonated 15N-, 13C-, 2H-labeled proteins. J. Biomol. NMR *13*, 369–374.

Gross, S., and Moore, C.L. (2001a). Rna15 interaction with the A-rich yeast polyadenylation signal is an essential step in mRNA 3′-end formation. Mol. Cell. Biol. *21*, 8045–8055.

Gross, S., and Moore, C. (2001b). Five subunits are required for reconstitution of the cleavage and polyadenylation activities of *Saccharomyces cerevisiae* cleavage factor I. Proc. Natl. Acad. Sci. USA *98*, 6080–6085.

Hammell, C.M., Gross, S., Zenklusen, D., Heath, C.V., Stutz, F., Moore, C., and Cole, C.N. (2002). Coupling of termination, 3′ processing and mRNA export. Mol. Cell. Biol. *22*, 6441–6457.

Hockert, J.A., Yeh, H.J., and MacDonald, C.C. (2010). The hinge domain of the cleavage stimulation factor protein CstF-64 is essential for CstF-77 interaction, nuclear localization, and polyadenylation. J. Biol. Chem. *285*, 695–704.

Hubbard, S.J., and Thornton, J.M. (1993). NACCESS, Computer Program (London: Department of Biochemistry and Molecular Biology, University College).

Johnson, B.A., and Blevins, R.A. (1994). NMRView: a computer program for the visualization and analysis of NMR data. J. Biomol. NMR *4*, 603–614.

Kessler, M.M., Zhao, J., and Moore, C.L. (1996). Purification of the *Saccharomyces cerevisiae* cleavage/polyadenylation factor I. J. Biol. Chem. *271*, 27167–27175.

Kim, M., Ahn, S.-H., Krogan, N.J., Greenblatt, J.F., and Buratowski, S. (2004). Transitions in RNA polymerase II elongation complexes at the 3′ ends of genes. EMBO J. *23*, 354–364.

Laskowski, R.A., Rullmannn, J.A., MacArthur, M.W., Kaptein, R., and Thornton, J. (1996). AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. J. Biomol. NMR *8*, 477–486.

Leeper, T.C., Qu, X., Lu, C., Moore, C., and Varani, G. (2010). Novel protein-protein contacts facilitate mRNA 3′-processing signal recognition by Rna15 and Hrp1. J. Mol. Biol. *401*, 334–349.

Legrand, P., Pinaud, N., Minvielle-Sebastia, L., and Fribourg, S. (2007). The structure of the CstF-77 homodimer provides insight into CstF assembly. Nucleic Acids Res. *35*, 4515–4522.

Lutz, C.S. (2008). Alternative polyadenylation: a twist on mRNA 3′ end formation. ACS Chem. Biol. *3*, 609–617.

Mackereth, C.D. (2011). Chemical shift assignments of a minimal Rna14p/Rna15p heterodimer from the yeast cleavage factor IA complex. Biomol. NMR Assign. *5*, 93–95.

Madl, T., Bermel, W., and Zangger, K. (2009). Use of relaxation enhancements in a paramagnetic environment for the structure determination of proteins using NMR spectroscopy. Angew. Chem. Int. Ed. Engl. *48*, 8259–8262.

Mandart, E., and Parker, R. (1995). Effects of mutations in the *Saccharomyces cerevisiae RNA14, RNA15,* and *PAP1* genes on polyadenylation in vivo. Mol. Cell. Biol. *15*, 6979–6986.

Mandel, C.R., Bai, Y., and Tong, L. (2008). Protein factors in pre-mRNA 3′-end processing. Cell. Mol. Life Sci. *65*, 1099–1122.

Millevoi, S., and Vagner, S. (2010). Molecular mechanisms of eukaryotic pre-mRNA 3′ end processing regulation. Nucleic Acids Res. *38*, 2757–2774.

Minvielle-Sebastia, L., Winsor, B., Bonneaud, N., and Lacroute, F. (1991). Mutations in the yeast *RNA14* and *RNA15* genes result in an abnormal mRNA decay rate; sequence analysis reveals an RNA-binding domain in the RNA15 protein. Mol. Cell. Biol. *11*, 3075–3087.

Minvielle-Sebastia, L., Preker, P.J., and Keller, W. (1994). RNA14 and RNA15 proteins as components of a yeast pre-mRNA 3′-end processing factor. Science *266*, 1702–1705.

Mulder, F.A. (2009). Leucine side-chain conformation and dynamics in proteins from 13C NMR chemical shifts. ChemBioChem *10*, 1477–1479.

Nedea, E., He, X., Kim, M., Pootoolal, J., Zhong, G., Canadien, V., Hughes, T., Buratowski, S., Moore, C.L., and Greenblatt, J. (2003). Organization and function of APT, a subcomplex of the yeast cleavage and polyadenylation factor involved in the formation of mRNA and small nucleolar RNA 3′-ends. J. Biol. Chem. *278*, 33000–33010.

Nilges, M. (1995). Calculation of protein structures with ambiguous distance restraints. Automated assignment of ambiguous NOE cross peaks and disulphide connectivities. J. Mol. Biol. *245*, 645–660.

Noble, C.G., Walker, P.A., Calder, L.J., and Taylor, I.A. (2004). Rna14-Rna15 assembly mediates the RNA-binding capability of *Saccharomyces cerevisiae* cleavage factor IA. Nucleic Acids Res. *32*, 3364–3375.

Pancevac, C., Goldstone, D.C., Ramos, A., and Taylor, I.A. (2010). Structure of the Rna15 RRM-RNA complex reveals the molecular basis of GU specificity in transcriptional 3′-end processing factors. Nucleic Acids Res. *38*, 3119–3132.

Puig, O., Caspary, F., Rigaut, G., Rutz, B., Bouveret, E., Bragado-Nilsson, E., Wilm, M., and Séraphin, B. (2001). The tandem affinity purification (TAP) method: a general procedure of protein complex purification. Methods *24*, 218–229.

Qu, X., Perez-Canadillas, J.-M., Agrawal, S., De Baecke, J., Cheng, H., Varani, G., and Moore, C. (2007). The C-terminal domains of vertebrate CstF-64 and its yeast orthologue Rna15 form a new structure critical for mRNA 3′-end processing. J. Biol. Chem. *282*, 2101–2115.

Preker, P.J., and Keller, W. (1998). The HAT helix, a repetitive motif implicated in RNA processing. Trends Biochem. Sci. *23*, 15–16.

Respondek, M., Madl, T., Göbl, C., Golser, R., and Zangger, K. (2007). Positioning of micelle-bound peptides by paramagnetic relaxation enhancements. J. Am. Chem. Soc. *129*, 5228–5234.

Rigaut, G., Shevchenko, A., Rutz, B., Wilm, M., Mann, M., and Seraphin, B. (1999). A generic protein purification method for protein complex characterization and proteome exploration. Nat. Biotechnol. *17*, 1030–1032.

Romier, C., Ben Jelloul, M., Albeck, S., Buchwald, G., Busso, D., Celie, P.H., Christodoulou, E., De Marco, V., van Gerwen, S., Knipscheer, P., et al. (2006). Co-expression of protein complexes in prokaryotic and eukaryotic hosts: experimental procedures, database tracking and case studies. Acta Crystallogr. D Biol. Crystallogr. *62*, 1232–1242.

Rouillard, J.M., Brendolise, C., and Lacroute, F. (2000). Rna14p, a component of the yeast nuclear cleavage/polyadenylation factor I, is also localised in mitochondria. Mol. Gen. Genet. *262*, 1103–1112.

Ruepp, M.D., Schweingruber, C., Kleinschmidt, N., and Schümperli, D. (2011). Interactions of CstF-64, CstF-77 and symplekin: implications on localization and function. Mol. Biol. Cell *22*, 91–104.

Shen, Y., Delaglio, F., Cornilescu, G., and Bax, A. (2009). TALOS+: a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. J. Biomol. NMR *44*, 213–223.

Simonelig, M., Elliott, K., Mitchelson, A., and O'Hare, K. (1996). Interallelic complementation at the suppressor of forked locus of Drosophila reveals complementation between suppressor of forked proteins mutated in different regions. Genetics *142*, 1225–1235.

Skaar, D.A., and Greenleaf, A.L. (2002). The RNA polymerase II CTD kinase CTDK-I affects pre-mRNA 3′ cleavage/polyadenylation through the processing component Pti1p. Mol. Cell *10*, 1429–1439.

Torchet, C., Bousquet-Antonelli, C., Milligan, L., Thompson, E., Kufel, J., and Tollervey, D. (2002). Processing of 3′-extended read-through transcripts by the exosome can generate functional mRNAs. Mol. Cell *9*, 1285–1296.

Worth, C.L., and Blundell, T.L. (2009). Satisfaction of hydrogen-bonding potential influences the conservation of polar sidechains. Proteins *75*, 413–429.

Wright, P.E., and Dyson, H.J. (2009). Linking folding and binding. Curr. Opin. Struct. Biol. *19*, 31–38.